

Hume on Mental Transparency

Hsueh Qu

Abstract: This paper investigates Hume's account of mental transparency. In this paper, I will endorse Qualitative Transparency – that is, the thesis that we cannot fail to apprehend the qualitative characters of our current perceptions, and these apprehensions cannot fail to be veridical – on the basis that, unlike its competitors, it is both weak enough to accommodate the introspective mistakes that Hume recognises, and yet strong enough to make sense of his positive employments of mental transparency. Moreover, Qualitative Transparency is also philosophically satisfying in providing good philosophical reason for why the mental states that are incorrigible should in fact be so.

1. A Map of the Literature

Mental transparency comprises two theses: incorrigibility, which is the thesis that we cannot be mistaken with respect to (at least some aspects of) our mental states;¹ and luminosity, which is the thesis that we cannot fail to be aware of (at least some aspects of) our mental states.² Hume seems to commit himself to transparency in the following passages:³

...all sensations are felt by the mind, such as they really are... (THN 1.4.2.5; SBN 189)⁴

[N]or is it conceivable that our senses shou'd be more capable of deceiving us in the situation and relations, than in the nature of our impressions. For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different. This were to suppose, that even where we are most intimately conscious, we might be mistaken. (THN 1.4.2.7; SBN 190)

But consciousness never deceives. (EHU 7.13; SBN 66)

...the perceptions of the mind are perfectly known.... (THN 2.2.6.2; SBN 366)

Hume scholars have often interpreted Hume as committed only to incorrigibility.⁵ Four interpretations of Hume on incorrigibility have been suggested in the literature:⁶

Total Incorrigibility: We cannot be mistaken about any aspect of our perceptions.⁷

Careful Incorrigibility: We cannot be mistaken about any aspect of our perceptions, so long as they are carefully considered.

Apperceptual Incorrigibility: Our higher-order perceptions of our perceptions as perceptions cannot be mistaken.

Qualitative Incorrigibility: Our apprehensions of the qualitative characters of our current perceptions cannot fail to be veridical.

Total Incorrigibility seems to be advocated by Stroud (1977), when he argues that Hume violates incorrigibility in maintaining that we can mistake calm passions for the actions of reason (THN 2.3.3.8; SBN 417). Careful Incorrigibility is hinted at by Baillie (2000) when he defends Hume's appeal to the calm passions by suggesting that a refined observer may be able to discern a calm passion amidst more violent emotions. Apperceptual Incorrigibility is defended by Baxter (2008, Chapter 5; and 2009), who argues that this reading makes sense of Hume's ruminations regarding personal identity in the Appendix. Qualitative Incorrigibility is endorsed in some form by Bricke (1980, Chapter 7) and Garrett (2009), although they phrase the principle in terms of 'mere' (that is, pre-conceptual; prior to our imaginative response to perceptions in classifying them using abstract ideas) and 'current' awareness respectively. These formulations grasp at the same concept, as the 'mere' or 'current' awareness of a perception simply consists in its phenomenal feel. Here I take 'qualitative character' to be the raw phenomenal feel or 'what-it-is-like-ness' of a perception.

I will ultimately endorse Qualitative Incorrigibility, and argue that this position entails Qualitative Luminosity, the thesis that we cannot fail to apprehend the qualitative characters of our current perceptions. In short, I read Hume as endorsing Qualitative Transparency:

Qualitative Transparency: We cannot fail to apprehend the qualitative characters of our current perceptions, and these apprehensions cannot fail to be veridical.

My paper will proceed as follows. In the following three sections, I examine Total, Careful, and Apperceptual Incorrighibility, arguing that they suffer from serious problems. Then, in Section 5, I examine Qualitative Incorrighibility, and find it to be both textually substantiated and philosophically tenable; I also argue that this position entails Qualitative Luminosity, and thus Hume should be read as endorsing Qualitative Transparency. In Section 6, I make use of Qualitative Transparency to make sense of Hume's often opaque and convoluted arguments invoking mental transparency in THN 1.4.2.7 (ABN 190), EHU 7.13 (SBN 66), and THN 2.2.6.2 (SBN 366); I then conclude in Section 7 with a comparison to Berkeley.

2. Total Incorrighibility

Total Incorrighibility fails to make sense of Hume's frequent recognitions of our fallibility with respect to mental states. Some prominent examples:

Classificatory Mistakes: Habitual liars mistaking an idea of the imagination for a memory (THN 1.3.5.6; SBN 86); mistaking ideas for impressions, as in the case of 'sleep' and 'fever' (THN 1.1.1.1; SBN 2) or 'madness or folly' (THN 1.3.10.9; SBN 123); mistaking the calm passions for the actions of reason (THN 2.3.3.8; SBN 417).

Mistakes about Causal Relations: Mistaking internal impressions for external ones, e.g. the impressions of necessary connection, sounds, and smells (THN 1.3.14.25; SBN 167); mistaking the effective motive for an action (EPM App 2.7; SBN 299).

Mistakes about Intentional Objects: Mistaking a single desire for 999 pounds for a single desire for a thousand pounds (THN 1.3.12.24; SBN 141); mistaking a thousand desires for a pound for a single desire for a thousand pounds (*ibid.*); mistaking resembling perceptions of temporally unextended impressions for a perception of a single temporally extended object (THN 1.4.2.40; SBN 207-8); mistaking a bundle of perceptions for an idea of a single self (THN 1.4.6.5; SBN 253).

To read Hume as being so explicitly and repeatedly inconsistent would be highly uncharitable. Moreover, it is difficult to provide any account or explanation of why Total Incorrighibility should hold. Total Incorrighibility thus seems unsatisfactory.

3. Careful Incorrighibility

Careful Incorrighibility on the other hand allows for various forms of error regarding our mental states, avoiding most purported counterexamples by blaming carelessness or inattention for such mistakes. This is a natural move to make, paralleling Hume's Conceivability Principle, which states that conceivability implies possibility (THN 1.1.7.6; SBN 19-20, THN 1.2.2.8; SBN 32); this principle is not falsified by any fool who thinks himself to conceive valley-less mountains, because it includes a caveat that the ideas we conceive of must be *clear and distinct* ones. Likewise, it may be claimed that any incorrighibility principle should possess a similar caveat, merely making the more modest claim that we cannot be mistaken regarding those perceptions that are carefully examined.

However, this reading seems implausible given that in Hume's characterisations of mental transparency, Hume emphasises the *impossibility* of mistake. Examine: '**nor is it conceivable** that our senses shou'd be more capable of deceiving us... actions and sensations of the mind... must **necessarily** appear in every particular what they are, and be what they appear... 'tis **impossible** any thing shou'd to *feeling* appear different' (THN 1.4.2.7; SBN 190, boldface added); 'consciousness **never** deceives' (EHU 7.13; SBN 66, boldface added); 'the perceptions of the mind are **perfectly** known' (THN 2.2.6.2; SBN 366, boldface added). It is thus doubtful that Hume thinks incorrighibility holds only when we are being careful.

Another problem is Hume's claim in THN 1.2.4.19 (SBN 45) that 'the same number of minima' is not a usable standard of equality, as minima are 'confounded with each other' and it is 'utterly impossible' for the mind to compute their number.⁸ Whatever judgment we made on 'the number of minima' contained in a spatial perception could never be incorrighible as no amount of care could overcome this difficulty. Moreover, as with Total Incorrighibility, it is

difficult to explain why judgments about carefully examined perceptions should be incorrigible – it is unclear why the ‘carefulness’ caveat should give us *incorrigibility* rather than mere reliability. Careful Incorrigibility thus also seems an unsatisfactory interpretation.

4. Apperceptual Incorrigibility

Baxter (2008) reads Hume’s remark that ‘Most philosophers seem inclin’d to think, that personal identity *arises* from consciousness; and consciousness is nothing but a reflected thought or perception’ (THN App 20; SBN 635) as indicating that consciousness consists in perceptions of perceptions as perceptions; taken in conjunction with Hume’s claim that ‘consciousness never deceives’ (EHU 7.13; SBN 66), this means that perceptions of perceptions as perceptions cannot be mistaken.⁹ However, Garrett (2009) disagrees with Baxter’s reading, arguing that the phrase ‘consciousness is nothing but a reflected thought or perception’ falls under the scope of what Hume says ‘most philosophers seem inclin’d to think’, rather than reflecting his own opinion.¹⁰

However, even if we grant that Hume thinks consciousness to be ‘nothing but a reflected thought or perception’, Baxter’s reading of ‘reflected thought or perception’ as ‘perception of a perception as a perception’ is nevertheless problematic.¹¹ A ‘perception of a perception as a perception’ is a perception that has as its intentional object another perception and recognises its object as a perception (rather than as an object external to the mind). But this is an implausible reading of what Hume means by ‘reflected thought or perception’ in context. In saying that ‘Most philosophers seem inclin’d to think, that personal identity *arises* from consciousness; and consciousness is nothing but a reflected thought or perception’ (THN App 20; SBN 635), Hume is clearly referencing Locke’s treatment of personal identity, which takes personal identity to turn on memory.¹² In short, ‘reflected thought or perception’ encompasses memory in this context. However, it is doubtful that Locke in particular would have thought that memory was a perception of a perception as a perception, since this would misidentify the intentional object of the memory: my memory of my mother is not directed at my impression

of my mother, but at my mother herself!¹³ Moreover, even if the intentional object of memories was taken to be impressions, memories would not take these impressions *as perceptions*, since memories do not typically engender an awareness of the status of the copied impression as a perception; my memory of my mother does not bring to mind any explicit considerations of perceptions in my mind. In short, ‘reflected thought or perception’ seems unlikely to have meant ‘perception of a perception as a perception’, which is the result that Baxter requires.

Moreover, it is unlikely that Hume uses ‘consciousness’ in a strict sense as indicating higher-order perception, given his usage of the term in a myriad of contexts. For instance:

... even with relation to that succession, we cou’d only admit of those perceptions, which are immediately present to our consciousness, nor cou’d those lively images, with which the memory presents us, be ever receiv’d as true pictures of past perceptions. (THN 1.4.7.3; SBN 265)

In distinguishing the perceptions with which ‘the memory presents us’ from those that are ‘immediately present to our consciousness’, the relevant distinction is between memories and *current* perceptions (rather than *higher-order* perceptions), which casts doubt on Baxter’s higher-order reading of ‘consciousness’. Although Hume might characterise consciousness as ‘a reflected thought or perception’ in the Appendix, this does not translate to a strict usage of this term, as he might merely be highlighting a connotation of consciousness (or some element of truth in other philosophers’ definitions of it) that is relevant to his purposes. Indeed, given the many ways in which Hume uses the term ‘consciousness’, he likely uses the term ambiguously, much like Clarke (1731), who gives *six* definitions of the term:¹⁴

"[E]ither the Reflex Act, by which a Man knows his Thoughts to be his own Thoughts; (*which is the strict and properest Sense of the Word;*) or the Direct Act of Thinking; or the Power or Capacity of Thinking; or (*which is of the same Import,*) simple Sensation; or the Power of Self-Motion, or of beginning Motion by the Will." (177n)

Notably, some of these definitions are immediate rather than higher-order, such as ‘the direct act of thinking’ and ‘simple sensation’. Reid also defines consciousness as immediate in the *Essays on the Intellectual Powers of Man*: ‘Consciousness is a word used by Philosophers, to

signify that immediate knowledge which we have of our present thoughts and purposes, and, in general, of all the present operations of our mind' (Essay 1 Chapter 1).¹⁵

And indeed, context renders it unlikely that Hume uses 'consciousness' to indicate higher-order perception in characterising mental transparency in EHU 7.13 (SBN 66) and THN 1.4.2.7 (SBN 190). Examine two passages that immediately follow Hume's claim that 'consciousness never deceives' in EHU 7.13 (SBN 66):

But if the original power were *felt*, it must be known. (EHU 7.14; SBN 66, emphasis added)

We may, therefore, conclude... that our idea of power is not copied from *any sentiment or consciousness* of power within ourselves... (EHU 7.15; SBN 67, emphasis added)

The above emphasis on feeling and sentiment suggests that Hume's point in claiming that 'consciousness never deceives' concerns *first-order* feeling or sentiment rather than *higher-order* perception. Moreover, in EHU 7.14 (SBN 66), Hume refers to the impression of power not 'being directly and fully known by an inward sentiment or consciousness', instead being 'mysterious and unintelligible' (EHU 7.14). Clearly, Hume thinks that had the impression of power been known by consciousness, it would be 'directly and fully known'; that is to say, consciousness involves *direct* access to a perception. This contrasts with Baxter's reading of consciousness as involving indirect higher-order perceptions of perceptions. Similarly, it is unlikely that Hume's claim that 'all actions and sensations of the mind are known to us by consciousness' in THN 1.4.2.7 (SBN 190) concerns higher-order perception given that it is invoked to guarantee the non-deceitfulness of the *senses*, which concern first-order impressions. This directly renders it doubtful that Hume's treatment of mental transparency concerns higher-order perception.

In any case, Baxter argues that Apperceptual Incorrigeability makes sense of Hume's second thoughts regarding his account of personal identity in the Appendix (THN App. 10-21; SBN 633-636). According to Baxter, Hume thinks his account of personal identity to be problematic because it claims that we confuse a perception of a bundle of perceptions for a

perception of a single self, thus contradicting Apperceptual Incorrigeability. For Baxter's interpretation to be viable, this problem should be unique to personal identity, since Hume frames his worry specifically to it. To ensure uniqueness, Baxter adds the stipulation that we must perceive our perceptions *as perceptions* for them to be incorrigible, crucially allowing for our mistaking perceptions for external objects.

Despite this, I believe uniqueness fails to hold. In mistaking a perception of a bundle of perceptions for a perception of a single self, our error (at least partly) concerns the intentional content of our perception of a bundle of perceptions, as we mistakenly think our perception to be *of a single self* rather than *of a bundle of perceptions*. However, as noted in Section 2, Hume recognises other types of mistakes about intentional content besides the cases of personal identity and the external world. For example, he takes us able to mistake a desire for 999 pounds for a desire for a thousand pounds (THN 1.3.12.24; SBN 141): our higher-order perception of a desire for 999 pounds would be a perception of a perception as a perception, but Hume believes us able to mistakenly take the desire to be for a thousand pounds instead. In the same passage, Hume also holds that we can mistake a thousand desires for a pound for a single desire for a thousand pounds (*ibid.*); we pre-philosophically treat the compound desire as a simple one, and so our perception of this desire (as a perception) is mistaken. Furthermore, this type of mistake is relevantly analogous to the case of personal identity in mistaking a perception of a bundle of perceptions for a perception of a single entity, and it seems difficult to draw some pertinent distinction between these two cases.¹⁶

Perhaps Apperceptual Incorrigeability could be defended by construing it to only extend to perceptions of *intrinsic* qualities of perceptions. Intentionality is an 'extraneous denomination' (THN 1.1.7.6; SBN 20); similarly, causal and classificatory properties are likewise extrinsic.¹⁷ However, I think this plausible amendment unsatisfactory for Baxter's purposes, given that the cases of a thousand desires for a pound and personal identity seem relevantly analogous; even if the relevant mistake does not concern intentional content, it seems difficult to see what mistake could apply to personal identity but not the thousand desires.

An interlocutor might reply that this objection engages only with Baxter's particular account; perhaps Hume takes neither his account of personal identity nor the case of a thousand passions to violate mental incorrigibility, leaving open the possibility that Hume endorses 'intrinsic' Apperceptual Incorrigibility. However, 'intrinsic' Apperceptual Incorrigibility seems to reduce to a form of Qualitative Incorrigibility, as the intrinsic qualities of a perception seem limited to its qualitative character; ideas shorn of their extrinsic properties seem merely experiential. I will argue in the next section that such a version of Qualitative Incorrigibility proves problematic in light of Hume's psychological atomism.

5. Qualitative Incorrigibility and the Appearance/Reality Gap

As noted earlier, 'qualitative character' indicates the phenomenal feel or 'what-it-is-like-ness' of a perception, which corresponds to what Hume calls its 'sensation' or 'feeling'.¹⁸

The *second* quality, which I discover in these passions, and which I likewise consider as an original quality, is their sensations, or the peculiar emotions they excite in the soul, and which constitute their very being and essence. Thus pride is a pleasant sensation, and humility a painful; and upon the removal of the pleasure and pain, there is in reality no pride nor humility. Of this our very feeling convinces us; and beyond our feeling, 'tis here in vain to reason or dispute. (THN 2.1.5.4; SBN 286)

The sensations of pride and humility constitute 'their very being and essence', and 'upon the removal of the pleasure and pain, there is in reality no pride and humility'. In short, these passions are wholly constituted by their sensations. Hume does note that another quality of pride and humility is their intentionality (THN 2.1.5.3; SBN 286), but intentionality is an extrinsic property: Hume notes that the intentionality of ideas is an 'extraneous denomination', 'of which in itself [an idea] bears no mark or character' (THN 1.1.7.6; SBN 20).¹⁹ Similarly, vivacity also seems to be an extrinsic feature of perceptions: Hume notes that '[a] particular shade of any colour may acquire a new degree of liveliness or brightness without any other variation. But when you produce any other variation, 'tis no longer the same shade or colour' (THN 1.3.7.5; SBN 96), which indicates that one can vary vivacity without changing the colour

perception in question. While qualitative character constitutes the ‘being and essence’ of pride and humility, vivacity surely does not, as one and the same passion of, say, pride at winning a pie-eating competition, could presumably lose vivacity (e.g. by considering the triviality of this achievement) while remaining the same passion.²⁰

This seems to generalise to other perceptions, given Hume’s imagistic psychology. Visual perceptions, for instance, are wholly constituted by visual experiences. A complex, visually extended idea of a book is composed of simple visual ideas, that is, *visual minima* (c.f. Garrett 1997, p. 53). A *visual minima* is clearly purely qualitative in nature, being the experience of a visual point. Similarly, an auditory perception seems wholly experiential. Likewise with abstract ideas: an abstract idea of a horse is a particular idea of a horse (a visual perception that is wholly qualitative) that has a disposition to revive other resembling perceptions via custom (THN 1.1.7.7; SBN 20). Intrinsically, perceptions are wholly constituted by their qualitative characters; there is nothing to a perception beyond its qualitative character. A perception is simply a single existent, without lateral content.²¹

With this in mind, we can distinguish two forms of Qualitative Incorrighibility:

Weak Qualitative Incorrighibility: Our apprehension of the qualitative characters of our current perceptions cannot fail to be veridical.²²

Strong Qualitative Incorrighibility: Our second-order beliefs about the qualitative characters of our current perceptions cannot be false.²³

Strong Qualitative Incorrighibility was examined in the previous section as an amendment of Apperceptual Incorrighibility. Like the previous interpretations examined, Strong Qualitative Incorrighibility struggles to provide an adequate explanation for why incorrighibility should hold. Armstrong (1963, p. 422) criticises mental incorrighibility on the grounds that ‘apprehension of the [current mental] occurrence will have to be *distinct* from the occurrence that is apprehended’; Mackie (1963, p. 23) raises the same objection. If these two elements are distinct, it is possible for them to differ, thus opening up the (at the very least metaphysical) possibility of error. A belief about a qualitative character is distinct from the qualitative

character itself; such beliefs therefore cannot be incorrigible because of this appearance/reality gap.

Contemporary defenders of versions of Strong Qualitative Incorrigibility generally address this objection by positing a metaphysically necessary connection (such as constitution or containment) between such beliefs and the qualitative characters they are about in order to close this appearance/reality gap. Although this move seems philosophically sound, it would be incompatible with Hume's atomistic psychology. Hume takes every perception to be a distinct, separate entity, with no necessary relations to anything else:

... all our perceptions are different from each other, and from every thing else in the universe, they are also distinct and separable, and may be consider'd as separately existent, and may exist separately, and have no need of any thing else to support their existence. (THN 1.4.5.5; SBN 233)

This thesis precludes a belief X about the qualitative character of, say, a particular impression of pride Y from having a necessary connection to its intentional object, since such a necessary connection would entail a necessary connection between this belief and the corresponding passion of pride: X would be necessarily connected to the qualitative character of Y, which is its 'very being and essence' (THN 2.1.5.4; SBN 286); consequently, X would be necessarily related to Y, which would violate Hume's psychological atomism.²⁴

And indeed, Hume recognises that higher-order beliefs about our mental states are corrigible: '... the operations of the mind... though most intimately present to us, yet, whenever they become the object of reflection, they seem involved in obscurity...' (EHU 1.13; SBN 13); and again: 'But the finer sentiments of the mind, the operations of the understanding, the various agitations of the passions, though really in themselves distinct, easily escape us, when surveyed by reflection...' (EHU 7.1; SBN 60).²⁵

Evidently, overcoming the 'Distinct Existences' objection is particularly difficult given Hume's psychological atomism. The only way to circumvent this objection is for both the 'grasping' (for want of a better term) and the thing grasped to be *exactly identical*. In

characterising Hume on mental transparency, we must select suitable notions of ‘grasping’ and ‘thing grasped’ so as to ensure this identity.

An important lesson gleaned from our examination of Strong Qualitative Incorrigeability is that it is mistaken in its characterisation of ‘grasping’. Where ‘grasping’ takes the form of a further idea or impression of the thing grasped (such as a belief regarding the thing grasped), the two must be distinct. However, Hume has a rather sparse mental ontology; if ‘grasping’ a mental state or aspect thereof does not involve forming a further perception of it, then what mental aspect does it involve? The only available answer is that the relevant form of ‘grasping’ is simply the *experience itself* of the mental state or aspect thereof. Let us call this form of ‘grasping’ *apprehension*: to apprehend X is to experience X.

Given this notion of grasping, what then is a suitable candidate for the thing apprehended that will close the appearance/reality gap? The answer has to itself be an experience, since apprehending X just is experiencing X. Upon reflection, it seems clear that the thing apprehended has to be qualitative character, that is, the ‘sensation’ of a perception, which just is exactly an experience itself. The experience of a qualitative character is exactly identical to the qualitative character itself; thus, there is no appearance/reality gap. To apprehend a qualitative character is merely to experience it; the apprehension of a qualitative character is not a further action performed over and above having the qualitative character occur in one’s mind. Putting this together, the relevant version of mental incorrigibility we derive is Weak Qualitative Incorrigeability: our apprehensions of qualitative characters cannot fail to be veridical. Weak Qualitative Incorrigeability thus avoids the ‘Distinct Existences’ objection and thereby accounts for the necessity of incorrigibility.²⁶

We are now in a position to see that Qualitative Incorrigeability²⁷ entails Qualitative Luminosity, that is, the thesis that we always apprehend our qualitative characters. Since there is no distinction between a qualitative character and our apprehension of it, if a qualitative character exists, necessarily, we apprehend it. The same fact that guarantees Qualitative

Incorrigibility also ensures Qualitative Luminosity.²⁸ Thus, Hume endorses Qualitative Transparency.

Qualitative Transparency is also textually substantiated by Hume's statements regarding mental transparency, as I will argue. Note that Qualitative Luminosity is not a thesis Hume incidentally commits himself to in virtue of his endorsing Qualitative Incorrigibility. In the passages endorsing mental transparency, we can see that Hume endorses both Qualitative Incorrigibility and Qualitative Luminosity. Examine:

Upon this head we may observe, that all sensations are felt by the mind, such as they really are, and that when we doubt, whether they present themselves as distinct objects, or as mere impressions, the difficulty is not concerning their nature, but concerning their relations and situation. (THN 1.4.2.5; SBN 189)

The claim that 'all sensations are felt by the mind' endorses Qualitative Luminosity given the reference to *feeling*, while the claim 'such as they really are' endorses Qualitative Incorrigibility. Moreover, the 'nature' of our sensations – that is, what they are intrinsically wholly constituted by – just is their qualitative character; the claim that this is transparent affirms Qualitative Transparency.

Similarly:

... all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different. (THN 1.4.2.7; SBN 190)

The claims that 'all actions and sensations of the mind are known to us by consciousness' and that they 'must appear in every particular what they are' clearly endorses luminosity, while the claim that they must 'be what they appear' concerns incorrigibility. That Hume is concerned with the transparency of qualitative character is substantiated by his emphasis on *feeling* in this passage.

Likewise:

...the perceptions of the mind are perfectly known.... (THN 2.2.6.2; SBN 366)

Here Hume is at a stroke endorsing both luminosity (because the perceptions of the mind are known) and incorrigibility (because they are *perfectly* known).

In the same vein, as previously argued in Section 4, Hume's emphasis on feeling and sentiment in the passages surrounding EHU 7.13 (SBN 66) suggests his statement that 'consciousness never deceives' there concerns qualitative character:

But if the original power were felt, it must be known. (EHU 7.14; SBN 66, emphasis added)

We may, therefore, conclude... that our idea of power is not copied from any sentiment or consciousness of power within ourselves... (EHU 7.15; SBN 67, emphasis added)

Note also that Hume's statement in EHU 7.14 (SBN 66) that the feeling of a power entails its being known, seems to be a statement invoking mental luminosity rather than incorrigibility. In sum, there is strong textual evidence that Hume endorses Qualitative Transparency.²⁹

Note that Qualitative Transparency is as capable as Apperceptual Incorrigibility in accommodating Hume's usage of 'consciousness' in characterising mental transparency in EHU 7.13 (SBN 66) and THN 1.4.2.7 (SBN 190). Consciousness has a clear connotation of qualitative character, as Block's (1980) 'China Brain' thought experiment establishes; similarly, Chalmers (1995) calls the question of how and why we have qualitative experience the 'hard problem of consciousness'. Neither is such a usage historically anomalous: one of Clarke's (1731, 1777n) six definitions of 'consciousness' is as 'simple Sensation', which is suggestive of qualitative character. Malabranche too seems to identify 'consciousness' and 'thought' with 'inner sensation', which also seems to indicate qualitative character:³⁰

[B]y the words *thought, mode of thinking, or modification of the soul*, I generally understand all those things that cannot be in the soul without the soul being aware of them through the inner sensation it has of itself. (*Search* 218)

It is true that we know well enough through our consciousness, or the inner sensation that we have of ourselves, that our soul is something of importance. (*Search* 238)

And indeed, Hume's usage of 'consciousness' can almost always be unproblematically read as having a connotation of qualitative character.³¹ All in all, it seems plausible that, in stating the thesis that 'consciousness never deceives', Hume affirms Qualitative Transparency.³²

In this section, I have hoped to establish that Qualitative Transparency is both philosophically tenable and also textually supported by the language Hume uses in alluding to mental transparency. In the next section, I show that besides these advantages, Qualitative Transparency allows us to clarify and make sense of Hume's rather opaque arguments invoking mental transparency in THN 1.4.2.7 (SBN 190), EHU 7.13 (SBN 66), and THN 2.2.6.2 (SBN 189).

6. Making Sense of Hume's Arguments

As defended above, Qualitative Transparency is a fairly weak thesis, since it guarantees only our apprehensions of qualitative characters. This allows it to unproblematically accommodate the various mistakes regarding our mental states that Hume allows for. The transparency of the apprehension of qualitative character, i.e. the bare experience of a perception, does not rule out classificatory, causal, and intentional mistakes regarding our perceptions, which all involve higher-order beliefs *about* our perceptions. While the weakness of Qualitative Transparency allows for these various forms of error, one worry is that this very weakness might entail that Qualitative Transparency proves inadequate in explaining Hume's positive employment of mental transparency in *ruling out* some forms of error. I argue that each of Hume's appeals to mental transparency substantiates Qualitative Transparency; however, the arguments lean more explicitly on incorrigibility, which might explain why some commentators have interpreted Hume as concerned only with incorrigibility.

a) The Impossibility of the Senses Deceiving Us

Hume's convoluted 'Of Scepticism with regard to the Senses' is concerned with the question of how we take sensations (viz. those impressions conveyed by the senses) to have continuous and distinct existence apart from ourselves. Hume appeals to Qualitative Transparency twice in arguing for the lemma that the senses are not the responsible party: in arguing that the senses do not deceive us with respect to the 'nature' (THN 1.4.2.5; SBN 189) and also the 'situation and relations' (THN 1.4.2.7; SBN 190) of our sensations.

Hume first rules out the senses convincing us of the 'continuous existence' of our sensations apart from ourselves, since to do so, the senses would have to operate when they cease to operate, which is 'a contradiction in terms' (THN 1.4.2.3; SBN 188). Hume next rules out the senses being responsible for our taking sensory impressions to have distinct existence from ourselves. Given that the senses convey only 'a single perception', and 'nothing beyond' (THN 1.4.2.4; SBN 189), if they are to convince us of the distinctness of our sensations from ourselves, it must be through 'fallacy and illusion' (THN 1.4.2.5; SBN 189) – that is, an element of deception must be involved. There are in principle two respects in which we could be deceived regarding our sensations: either with regard to their 'nature', or with regard to their 'relations and situations' (THN 1.4.2.5; SBN 189).

Hume argues the senses cannot deceive us regarding the 'nature' of our sensations (THN 1.4.2.5; SBN 189). Hume then dismisses the possibility of the senses deceiving us with respect to the 'situation and relations' of our sensations. Hume briefly makes the point in THN 1.4.2.5-6 (SBN 189-90) that the senses cannot distinguish between ourselves and external objects and thus cannot represent perceptions as being distinct from ourselves, before raising what he takes to be the decisive blow against thinking the senses deceive us in this respect:

Add to this, that every impression, external and internal, passions, affections, sensations, pains and pleasures, are originally on the same footing; and that whatever other differences we may observe among them, they appear, all of them, in their true colours, as impressions or perceptions. And indeed, if we consider the matter aright, 'tis scarce possible it shou'd be otherwise, nor is it

conceivable that our senses shou'd be more capable of deceiving us in the situation and relations, than in the nature of our impressions. For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different. This were to suppose, that even where we are most intimately conscious, we might be mistaken. (THN 1.4.2.7; SBN 190)

Here Hume argues for the conclusion that the senses are not 'capable of deceiving us in the situation and relations... of our impressions'. The key moves establishing this thesis are his claims that 'actions and sensations of the mind... must necessarily appear in every particular what they are, and be what they appear' and that 'every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different'. I interpret Hume's argument as leaning on Qualitative Incorrighibility in the following way:³³

1. Every perception's qualitative character is incorrighible. (Qualitative Incorrighibility)
 - 'Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different'.
2. Actions and sensations of the mind are impressions that are wholly constituted by qualitative character and moreover have no representational content. (Implicit)
3. If actions and sensations of the mind are wholly constituted by qualitative character and moreover have no representational content, and every perception's qualitative character is incorrighible, then actions and sensations of the mind cannot deceive us. (Implicit)
4. Therefore, actions and sensations of the mind cannot deceive us. (From 1, 2 and 3)
 - 'For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear'.
5. The senses are the faculty that solely convey sensations. (Implicit)
6. If the senses convey only perceptions that cannot deceive us, then the senses cannot deceive us. (Implicit)
7. Therefore, the senses do not deceive us. (From 4, 5, and 6)

8. If the senses do not deceive us, then they cannot deceive us with respect to the situation and relations of our impressions. (Implicit)
9. Therefore, the senses cannot deceive us with respect to the situation and relations of our impressions. (From 7 and 8)
 - '[N]or is it conceivable that our senses shou'd be more capable of deceiving us in the situation and relations, than in the nature of our impressions?'

One question that arises regarding premise 4 is this: what are the 'actions and sensations of the mind'? The sensations of the mind seem to be simply impressions of sensation. 'Actions of the mind' are more difficult to characterise, but an examination of the relevant passage illuminates the matter. Hume glides seamlessly from stating that 'every impression, external and internal, passions, affections, sensations, pains and pleasures' must 'appear... in their true colours' to speaking of 'actions and sensations of the mind' appearing in every particular what they are, and being what they appear; clearly, Hume thinks 'actions and sensations of the mind' to be coextensive with 'every impression, external and internal', and so 'actions of the mind' refers to those impressions that are not sensations (viz. internal impressions) in this context.³⁴

To establish his claim that actions and sensations of the mind cannot deceive us, Hume appeals to the more general thesis that is Qualitative Incorrigeability: 'Every thing that enters the mind, being in *reality* a perception, 'tis impossible any thing shou'd to *feeling* appear different', that is, the qualitative characters of all perceptions lack an appearance/reality gap, and must therefore be incorrigible.³⁵ This inference from Qualitative Incorrigeability to premise 4 proceeds on the basis of the assumption that impressions are wholly constituted by their qualitative characters, and moreover have no representational content. Hume explicitly notes that passions and volitions lack representational content, for example (THN 2.3.3.5; SBN 415, THN 3.1.1.9; SBN 458). Impressions of sensation are likewise plausibly non-representational, given that we have no notion of what these impressions would represent.³⁶ Representation requires both resemblance to (THN 1.4.5.3; SBN 233) and causal derivation from (THN 1.2.3.11; SBN 37) the represented entity; however Hume explicitly expresses not only his lack

of interest in the causes of our sensations, but also remarks on the impossibility of ever discovering them (THN 1.3.5.2; SBN 85, EHU 12.12; SBN 153). How could something be representative for us if we have no notion of what it is representing? Plausibly, Hume at the very least treats sensations as ‘for all intents and purposes’ non-representational.

Given that these impressions are wholly constituted by their qualitative characters and without representational content, and given that the qualitative characters of our perceptions are incorrigible, it follows that they cannot deceive us. This is why Hume claims that these impressions ‘must necessarily appear in every particular what they are, and be what they appear’; since these perceptions are wholly experiential and without representational content, there is no possibility of their deceiving us, either with respect to our apprehension of their qualitative characters or in what they purport to represent – they cannot appear to be anything that they are not, either through misapprehension or misrepresentation. Sensations are wholly constituted by qualitative characters, which, given Qualitative Incorrigibility, cannot be misapprehended. Moreover, sensations lack any extrinsic representational content, and so cannot misrepresent. Thus, sensations cannot deceive us either in virtue of our apprehension of them or in what they purport to represent, and so our sensations cannot deceive us.

Hume’s conclusion in THN 1.4.2.7 (SBN 190) is that the senses cannot deceive us with respect to even the ‘situations and relations’ of our impressions of sensation. To clarify how Hume arrives at this conclusion, it is worth asking what it would take for the senses to deceive us. The senses are defined functionally as the faculty that solely conveys sensations; thus, the only way for the senses to deceive us is by conveying sensations which deceive us.³⁷ Of course, as we just established, our sensations cannot deceive us; correspondingly, the senses cannot deceive us either. Given this, our senses cannot be the source of any error, whether with respect to the ‘situation and relations’ or the ‘nature’ of our impressions; in other words, ‘nor it is conceivable that our senses shou’d be more capable of deceiving us in the situation and relations, than in the nature of our impressions’ (THN 1.4.2.7; SBN 190).³⁸

b) The Palsy Passage

The logic of the argument of the Palsy Passage (EHU 7.13; SBN 66) according to which Qualitative Transparency plays a crucial role is not obvious. I argue that Hume invokes Qualitative Incorrigeability in arguing that we do not derive the idea of power or necessary connection³⁹ from the impression of volition. Let us examine the passage in its entirety:

A man, suddenly struck with a palsy in the leg or arm, or who had newly lost those members, frequently endeavours, at first, to move them, and employ them in their usual offices. Here he is as much conscious of power to command such limbs, as a man in perfect health is conscious of power to actuate any member which remains in its natural state and condition. But consciousness never deceives. Consequently, neither in the one case nor in the other, are we ever conscious of any power.

We learn the influences of our will from experience alone. (EHU 7.13; SBN 66)

Here Hume argues that the impression of volition cannot be the source of our idea of necessary connection because we can separate the impression of volition from the action(s) it produces – in other words, we can have an impression of causally inefficacious volition, and it cannot therefore be the source of our idea of necessary connection. Hume points to the case of the man with palsy (who has an inefficacious volition) to demonstrate this point; however, it is open to his opponents to respond that there are two different impressions of volition, one efficacious and one inefficacious, and it is the former that is the source of our idea of necessary connection. This is why Hume draws the comparison between the efficacious and inefficacious impressions of volition, pointing out that they are in fact the same impression, and hence the man is ‘as much conscious of power’ in both cases. Hume’s appeal to the man’s conscious experience reveals that Hume is providing a *phenomenological* basis for this similarity claim. And indeed, the phenomenology suggests that Hume is right to make this claim; however, it is still open to his opponents to respond that we are simply *mistaken* with respect to our conscious experience in this regard. It is at this point in the dialectic that Hume invokes mental incorrigibility to support the incorrigibility of comparisons between efficacious and inefficacious impressions of volition; this is substantiated by the fact that Hume appeals to mental transparency

immediately upon claiming that the man with palsy is ‘as much conscious of power’ as a man in perfect health.

Between what items is this comparison meant to hold? One component is clearly the current impression of inefficacious volition. But the other component cannot be an impression of efficacious volition had by a different ‘man in perfect health’, as surely no plausible reading of Hume on mental transparency says anything about the infallibility of our knowledge of perceptions in other minds. Clearly Hume has to mean by ‘a man in perfect health’ the stricken man in his former healthy state, which is why he stipulates the man to be ‘suddenly struck’ with palsy rather than afflicted from birth. Here Hume is comparing the man’s impression of inefficacious volition with his *memory* of efficacious volition and finding them to be the same, taking this discovery to be substantiated by mental incorrigibility.⁴⁰ This sameness can be established by guaranteeing that their qualitative characters are the same, since they are wholly constituted by qualitative character.

Qualitative Incorrigibility can justify the similarity claim by providing some measure of assurance regarding some of the three elements that lead to this conclusion:

- (1) The qualitative character of the current impression of volition;
- (2) The memory of the qualitative character of the past impression of volition; and
- (3) The exact resemblance between the two qualitative characters.

(3) is guaranteed by Hume’s theory of relations, as the two qualitative characters exactly resemble,⁴¹ and the relation of resemblance is a constant relation; this comparison of resemblance is therefore ‘intuitively certain’ (THN 1.3.3.2; SBN 179). This comparison is not incorrigible; like judgments of arithmetic (based on the constant relation of proportions of quantity and number) there is a possibility of mistake. Nevertheless, such intuitively certain judgments have a privileged epistemic status, and it is dialectically ineffective for Hume’s opponents to question them; we do not take such simple comparisons to be liable to error, in the same way that one would not typically challenge simple arithmetical judgments.

The incorrigibility of (1) comes for free, as it follows directly from Qualitative Incorrigibility. All that is left to ensure the correctness of the conclusion is incorrigibility with respect to (2), that is, the memory of the impression of volition. In a sense, Qualitative Incorrigibility does not *need* to guarantee (2) in order for us to make sense of Hume's appeal to it in the Palsy Passage; having already guaranteed (1), Hume's appeal to mental transparency would be justified. However, (2) is arguably the component that Hume's opponents would find most questionable. It would therefore be congenial to my interpretation if we could read Qualitative Incorrigibility as providing, if not a full guarantee, then at the very least some assurance regarding judgments about (2).

Qualitative Incorrigibility guarantees the qualitative character of the man's *memory* of volition; what is needed is some way to ensure that this qualitative character is identical with that of the original volition. Enter the Copy Principle:

...all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent. (THN 1.1.1.7; SBN 4)

Given simplicity, we can be sure that the qualitative character of the current memory is a faithful copy of some prior simple impression, since simple ideas are *exact* copies of the simple impressions that they derive from.⁴² Of course, the impression of volition is clearly simple, as Hume certainly recognises in considering it as a candidate for the impression of necessary connection in applying his Copy Principle, which entails only that *simple* ideas correspond to exactly resembling *simple* impressions. Given that we have an idea of the simple impression of volition, Hume's Copy Principle dictates that this impression must have been directly experienced, and moreover the idea and the impression must exactly resemble. Given further that this memory is qualitatively identical to the current impression of inefficacious volition that it is being compared to, we can therefore be sure that it is a faithful memory *of* an impression of volition; operating on the background assumption shared by the disputing parties that the prior impression of volition was efficacious (since it occurred prior to the onset of palsy), we can be assured that the qualitative character of the current memory (itself guaranteed

by Qualitative Incorrigoibility) is faithful to the qualitative character of some past efficacious volition, even if the palsied man misremembers the exact circumstances in which he had that impression. Therefore, the palsied man can be sure that the efficacious and inefficacious impressions of volition are identical.

This is not to say that Qualitative Incorrigoibility completely guarantees this judgment, as there remain possible sources of error. For example, the Copy Principle or the causal efficacy of the past impression might be doubted. However, even if Qualitative Incorrigoibility does not completely guarantee the palsied man's judgment, it manages to eliminate the more likely sources of error, thus conferring a measure of justification on this judgment. Furthermore, many of these remaining sources of error are taken as background assumptions by both parties in the debate, which makes it dialectically ineffective to question them. Hume's Copy Principle, for example, will largely be accepted by his Lockean opponents; similarly, his contention that the prior impression of volition was efficacious is uncontentious, since the disputants agree that the impression occurred before the onset of palsy. Qualitative Incorrigoibility does provide some assurance with respect to the controversial elements of the palsied man's judgment, particularly those that concern his memory of an efficacious impression of volition; therefore, Hume's appeal to it is dialectically effective.

c) Defending Hume's Philosophical Account of the Passions

Hume appeals to mental transparency to defend his account of the passions (specifically his account of the relation between love and hatred, benevolence and anger) from suspicions of error. He does so by contrasting such investigations with our knowledge of external bodies:

But as the perceptions of the mind are perfectly known, and I have us'd all imaginable caution in forming conclusions concerning them, I have always hop'd to keep clear of those contradictions, which have attended every other system. (THN 2.2.6.2; SBN 366))

One peculiarity of this passage is why, having stated that 'the perceptions of the mind are perfectly known', Hume feels the need to reassure his reader that he has used 'all imaginable

caution in forming conclusions concerning them'. Why does the former not render the latter redundant? This peculiarity is particularly easy for Qualitative Transparency to explain, given its modesty. The passions are wholly constituted by their qualitative character, and so Qualitative Transparency entails that Hume's apprehension of them must be veridical. In his theorising about the passions, Hume is forming second-order beliefs about these qualitative characters. However, given that the passions are simple impressions, the Copy Principle ensures that our ideas of their qualitative characters are perfect copies. However, Hume's account also involves the formation of beliefs regarding the 'situation and relations' of these passions (such as love's relation with benevolence, and hatred's relation with anger); these beliefs are *not* incorrigible, which is why he needs to make sure to use 'all imaginable caution in forming conclusions concerning them'. Nevertheless, given Qualitative Incorrigibility, Hume's account will enjoy a greater measure of epistemic security, as a prominent source of error is entirely ruled out; in contrast with investigations concerning external objects, our apprehension *and* ideas of the 'essence and composition' (THN 2.2.6.2; SBN 366) of the passions cannot be mistaken. Given the caution Hume treads with, he can have 'hop'd to keep clear of those contradictions, which have attended every other system' (*ibid.*).

7. Conclusion

We have seen that Qualitative Luminosity succeeds textually as an interpretation on three counts: first, it accommodates the various types of error Hume allows for with respect to mental states; second, it is textually substantiated by Hume's various statements of mental transparency; and third, it makes sense of Hume's positive employments of the principle in his often opaque arguments. In conjunction with the philosophical tenability of the thesis, we have good reason to attribute it to Hume. I end the paper by noting that in addition to the above strengths, Qualitative Luminosity also has historical precedent in Berkeley:

For since [our ideas] and every part of them exist only in the mind, it follows that there is nothing in them but what is perceived. (*Principles* 25)⁴³

Here Berkeley endorses mental transparency on the same basis Hume does. In claiming that there is nothing in our perceptions beyond what is perceived, Berkeley maintains the position I argued Hume holds: there is no secret component in a perception that we could be ignorant of or mistaken about; a perception is a single existent, without lateral content. Moreover, Berkeley also explicitly links transparency to qualitative character:

Colour, figure, motion, extension and the like, considered only as so many *sensations* in the mind, are perfectly known, there being nothing in them which is not perceived. (*Principles* 87)

Here Berkeley emphasises that various primary and secondary qualities are perfectly known (incorrigibility), and have nothing in them which is not perceived (luminosity), so long as they are *considered only as sensations in the mind*. In short, their qualitative characters are transparent *qua* feeling. It is not implausible to think that Hume's account of mental transparency, which so closely parallels Berkeley's, was perhaps inspired by it.⁴⁴

References

- Ainslie, Donald (2012). 'Reflections on Penelhum on Hume and Locke', *Proceedings of the 39th International Hume Society Conference*.
- Armstrong, David (1963). 'Is introspective knowledge incorrigible?', *Philosophical Review* 62: 417–32.
- Armstrong, David (1981). 'What is consciousness?', in *The Nature of Mind*. Ithaca, NY: Cornell University Press, 55-67.
- Baillie, James (2000). *Hume on Morality*, Routledge.
- Block, Ned (1980). 'Troubles with Functionalism', in *Readings in the Philosophy of Psychology*, Volume 1, Ned Block, ed., Cambridge, Mass: Harvard University Press, 268-305.
- Baxter, Donald (2008). *Hume's Difficulty: Time and Identity in the Treatise*, Routledge.

- Baxter, Donald (2009). 'Replies to Perry, Falkenstein, and Garrett', *Philosophical Studies* 146: 445-455.
- Berkeley, George (1710). *Principles of Human Knowledge*, edited by H. Robinson. Oxford University Press, 1996.
- Bricke, John (1980). *Hume's Philosophy of Mind*, Princeton University Press.
- Clarke, Samuel (1731). *A Letter to Mr. Dodwell*, 6th edn, edited by James and John Knapton.
- Chalmers, David (1995). 'Facing Up to the Problem of Consciousness', *Journal of Consciousness Studies* 2(3):200-19.
- Chalmers, David (2003). 'The content and epistemology of phenomenal belief', in *Consciousness: New philosophical perspectives*, Quentin Smith and Aleksandar Jokic (eds.), Oxford: Oxford, 220–272.
- Cohon, Rachel, and Owen, David (1997). 'Representation, Reason, and Motivation', *Manuscrito* 20:47-76.
- Fogelin, Robert (1984). 'Hume and the Missing Shade of Blue', *Philosophical and Phenomenological Research* 45:263-72.
- Garrett, Don (1997). *Cognition and Commitment in Hume's Philosophy*, Oxford/New York: Oxford University Press.
- Garrett, Don (2006). 'Hume's Naturalistic Theory of Representation', *Synthese* 152:301-319.
- Garrett, Don (2009). 'Difficult Times for Humean Identity?', *Philosophical Studies* 146:435-443.
- Gertler, Brie (2001). 'Introspecting phenomenal states', *Philosophy and Phenomenological Research*, 63: 305–328.
- Gustafsson, Johan (2010). 'Did Locke Defend the Memory Continuity Criterion of Personal Identity?', *Locke Studies* 10: 113-129.
- Hume, David (1739-1740). *A Treatise of Human Nature*, edited by L. A. Selby-Bigge, 2nd ed. revised by P.H. Nidditch, Oxford: Clarendon Press, 1975.

- Hume, David (1748). *Enquiries concerning Human Understanding and concerning the Principles of Morals*, edited by L. A. Selby-Bigge, 3rd edition revised by P. H. Nidditch, Oxford: Clarendon Press, 1975.
- Horgan, Terence, John L. Tienson, and George Graham (2006). 'Internal-world skepticism and mental self-presentation', in *Self-representational approaches to consciousness*, Uriah Kriegel and Kenneth Williford (eds.), Cambridge, MA: MIT, 191–207.
- Horgan, Terence, and Uriah Kriegel (2007). 'Phenomenal epistemology: What is consciousness that we may know it so well?', *Philosophical Issues*, 17(1): 123–144.
- Jackson, Frank (1973). 'Is There a Good Argument Against the Incorrigibility Thesis?', *Australasian Journal of Philosophy* 51: 51–62.
- Jorgenson, Larry (2010). 'Seventeenth-Century Theories of Consciousness', *Stanford Encyclopedia of Philosophy*.
- Locke, John (1700). *An Essay Concerning Human Understanding*, edited by P. H. Nidditch, New York: Oxford University Press, 1975.
- Lycan, William (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Mackie, J.L (1963). 'Are There Any Incorrigible Empirical Statements?', *Australasian Journal of Philosophy* 41 (1):12-28.
- Malabranche, Nicolas (1974-75). *The Search after Truth*, edited by T.M. Lennon, and P.J Olscamp. Cambridge University Press, 1997.
- Oxford English Dictionary* (1989). 2nd ed. OED Online. Oxford University Press. <<http://dictionary.oed.com/>>.
- Qu, Hsueh (Forthcoming). 'The Simple Duality: Humean Passions', *Canadian Journal of Philosophy, Supplementary Volume: Hume in Alberta*.
- Reid, Thomas (1785). *Essays on the Intellectual Powers of Man*, edited by William Hamilton, Edinburgh: Maclachlan and Stewart, 1863.
- Rosenthal, David (1993). 'Thinking that one thinks', in M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell, 197-223.
- Schmaltz, Tad (1996). *Malebranche's Theory of the Soul*, Oxford: Oxford University Press.

Stroud, Barry (1977). *Hume*, Routledge.

Waxman, Wayne (1994). *Hume's Theory of Consciousness*, Cambridge University Press.

Williamson, Timothy (2000). *Knowledge and its Limits*, Oxford University Press.

Winkler, Kenneth (1991). 'Locke on Personal Identity', *Journal of the History of Philosophy* 29(2): 201-226.

Department of Philosophy, National University of Singapore

¹ Note that I do not take mistake to be solely concerned with falsehood; notably, I take non-veridicality to count as a mistake – hallucinations might be said to be mistakes, for example.

² Some distinguish infallibility from incorrigibility: they take incorrigibility to mean the impossibility of our being corrected, and infallibility to mean the impossibility of our being mistaken. In this paper, I take incorrigibility to be synonymous with infallibility in meaning the impossibility of being mistaken, as do Armstrong (1963), Stroud (1977), and Chalmers (2003), among others.

³ It might be thought that some of Hume's claims cast doubt on his maintaining mental transparency, for instance his claim that 'the operations of the mind... Whenever they become the object of reflection... seem involved in obscurity' (EHU 1.13; SBN 13), and his claim that 'the finer sentiments of the mind, the operations of the understanding, the various agitations of the passions, though really in themselves distinct, easily escape us, when surveyed by reflection' (EHU 7.1; SBN 60). These passages can easily be addressed on my proposed interpretation of Hume on mental transparency, and I will do so later in the paper.

⁴ In references to Hume, 'THN' refers to the *Treatise of Human Nature*, 'EHU' to the *Enquiry Concerning Human Understanding*, and 'EPM' to the *Enquiry Concerning the Principles of Morals*. Arabic numerals refer to section and paragraph numbers (EHU and EPM), or to book, part, section, and paragraph numbers (THN). SBN numbers refer to pages in the Selby-Bigge and Nidditch editions of the *Treatise* and two *Enquiries*.

⁵ For ease of exposition, I refer to the phenomenon Hume discusses in these passages as 'mental transparency' without intending to beg the question of commentators who read Hume as committed only to incorrigibility.

⁶ I do not take this list to be exhaustive; later, in defending Qualitative Incorrigibility, I give reasons why it is preferable to any other account. I focus on interpretations that are defended or at least suggested in the Hume literature, and so this list may appear somewhat unsystematic.

⁷ In this paper I use 'perception' in the way Hume does, that is, to mean an idea or impression.

⁸ Thanks to Don Garrett for suggesting this objection. Note that I take mistaking the 'number of minima' in a perception to be primarily a classificatory mistake, in being a misapplication of a number-concept.

⁹ Baxter (2009) also offers the following passage in support of his interpretation: 'When we reflect on our past sentiments and affections, our thought is a faithful mirror, and copies its objects truly' (EHU 2.2; SBN 17-18). I am not convinced that Hume is here making a statement about the *incorrigibility* rather than the *reliability* of reflected thoughts and perceptions. Saying that 'our thought is a faithful mirror' seems to lack the emphasis on the impossibility of error present in his explicit statements of mental transparency, which I noted in the previous section.

¹⁰ There is some interesting debate between Garrett and Baxter on 18th century punctuation, but I will not go into the intricacies here.

¹¹ Thanks are owed to John Biro for suggesting this objection.

¹² The issue is of course a matter of debate in Locke studies, but I lack the space to argue for it here. Winkler (1991) ably defends the 'memory' interpretation on the basis of *Essay* 2.28.20, for instance, but see Gustafsson

(2010) for a dissenting view. References to Locke are to *An Essay Concerning Human Understanding*, edited by Nidditch, Oxford University Press (1975). Arabic numerals refer to book, chapter, and paragraph numbers.

¹³ As an anonymous referee notes, Baxter's view of the intentionality of memory is more plausible in Hume's case, given Hume's explicit scepticism about the external world (THN 1.4.2). Although as the same referee notes, Hume on the external world is a controversial interpretive matter, which perhaps lessens the force of this objection somewhat.

¹⁴ Donald Ainslie also argues that Hume's usage of 'consciousness' is ambiguous, in 'Reflections on Penelhum on Hume and Locke', presented at the 39th *International Hume Society Conference* in Calgary (2012); this talk brought the Clarke citation to my attention. References to Clarke are to *A Letter to Mr. Dodwell*, 6th edition, edited by James and John Knapton.

¹⁵ References to Reid refer to *Essays on the Intellectual Powers of Man*, edited by William Hamilton, Edinburgh: Maclachlan and Stewart, 1863.

¹⁶ An anonymous referee points out that there is a distinction between the two cases, namely that there could be something which was a desire for a thousand pounds, but there could not be something which was a unified, enduring, substantial self. I think that this disanalogy will not help the supporter of Apperceptual Incorrignibility, however, since even if there could have been a desire for a thousand pounds, there is in fact none, and so we are still making an error regarding a perception of a perception as a perception in this case. The only version of Apperceptual Incorrignibility that could exploit this disanalogy would be one that ruled out the possibility of error with respect to mistakenly taking a perception to be an impossible (in some sense of impossible) perception, which seems somewhat arbitrary and textually unmotivated.

¹⁷ Baxter (2008) hints at this approach when he claims that Apperceptual Incorrignibility is not threatened by the case of the calm passions: 'Hume is acknowledging difficulties with theorizing about perceptions – classifying them and explaining their causes and effects – not with the mere consciousness of them' (p. 73). It is unclear what the difference between 'mere consciousness' and 'theorizing' is supposed to amount to, but it might be characterised as the distinction between our discernment of intrinsic and extrinsic properties.

¹⁸ Note that 'sensation' is used by Hume to refer both to impressions of sensation, and to qualitative characters. In THN 2.1.5.4 (SBN 286), Hume means by 'sensations' the qualitative characters of the passions, rather than impressions of sensation. These passions are simple and admit of no constituent parts, and so Hume cannot be referring to constituent *impressions* in speaking of the 'sensations' of the passions. Similarly in THN 2.1.5.9 (SBN 288), THN 2.2.1.3 (SBN 330), THN 2.2.1.6 (SBN 331) and THN 2.2.2.3 (SBN 333).

¹⁹ One might distinguish the intentionality of our passions (determined by various causal and resemblance relations, such as the double relation of impressions and ideas in the case of the indirect passions) from the intentionality of other mental states, such as the intentionality of our ideas (also determined by causal and resemblance relations) and that of our sensations (determined by our imaginative responses to constant and coherent experience). Nevertheless, the point about intentionality being extrinsic seems to generalise; indeed, the double relation of impressions and ideas in the case of pride and humility is clearly extrinsic. I complicate this picture in my 'The Simple Duality: Humean Passions' (Forthcoming) by arguing that passions also have an intrinsic form of intentionality that is inseparable from and an aspect of their qualitative feel, but this does not unduly affect my arguments in this paper.

²⁰ Characterising 'force and vivacity' is a notoriously difficult task here, and I will not attempt to do so in this paper. Nevertheless, it seems that vivacity, being experiential in nature, plausibly seems as though it should be transparent on my account. Hume does think that we can make mistakes when comparing force and vivacity: for example, with regard to the case of confusing a desire for 999 pounds with a desire for a thousand, he says '... 'tis plainly impossible for the mind to run over every particular view, and distinguish the superior vivacity of the image arising from the superior number, when the difference is so inconsiderable' (THN 1.3.12.24; SBN 141). This mistake is partially one of intentional content (as noted in Sections 2 and 4), but also partially one of vivacity, as the latter desire would be marginally more vivacious than the former. However, one can easily construe these passages as describing classificatory mistakes about vivacity; we mistakenly classify a certain degree of vivacity as the same strength as a different degree of vivacity. Although I think vivacity is very plausibly transparent, I do not commit myself on this matter here, mainly because Hume's discussions of transparency do not involve vivacity, and so I take the issue to be somewhat textually underdetermined, particularly since Hume himself refrains from providing a detailed account of vivacity.

²¹ Thanks to Donald Ainslie for helping me to clarify my thoughts on the nature of qualitative character.

²² I use 'apprehension' here without its usual connotation of success to avoid triviality; I will more fully explicate this notion shortly. Note also that Weak Qualitative Incorrignibility does not involve truth or falsity, but rather veridicality and non-veridicality, as apprehensions are not truth-apt.

²³ The stronger version is more popular in contemporary literature: it is defended by Chalmers (2003); Horgan, Tienson, and Graham (2005); Horgan and Kriegel (2007); and Jackson (1973). Gertler (2001) shies away from infallibility, but takes second-order beliefs about qualitative character to possess ‘Cartesian certainty’.

²⁴ Logical connections can hold between the *contents* of beliefs, e.g. with some analytical truths. But this would not entail any metaphysically necessary connection between beliefs themselves, given that no analytic truths (or sets thereof) entail the existence of some belief such that it necessitated the existence of another.

²⁵ Of course, if Weak Qualitative Incorrigeability is true, beliefs about particular qualitative characters, although corrigible, would still possess more epistemic security than other beliefs in virtue of the privileged status of their content, as a source of possible error is entirely eliminated. Hume is being needlessly hyperbolic in these passages: in EHU 2.2 (SBN 17-18) he recognises that second-order beliefs about sensations and affections are ‘faithful mirror[s]’, for example.

²⁶ Note that in modern philosophy of mind, the position that apprehension of qualitative character is a direct awareness is defended by ‘same order’ theories of consciousness, which take qualitative characters to be self-presenting (c.f. Horgan, Tienson, and Graham [2005]; Horgan and Kriegel [2007]).

²⁷ Hereafter, I use ‘Qualitative Incorrigeability’ to mean the weaker version of this thesis, unless otherwise noted.

²⁸ Of course, Qualitative Luminosity does not preclude our forming false *beliefs* about these qualitative characters, or indeed failing to form any beliefs about them at all. Note that Williamson’s (2000, Chapter 4) well-known anti-luminosity argument does not impact Qualitative Luminosity, as Williamson’s argument applies to cases involving concepts such that one can move gradually from the concept obtaining to not obtaining (he gives the example of coldness). Apprehension is pre-conceptual, and so evades the argument.

²⁹ The interpretation defended here parallels Waxman (1994), who also defend mental transparency in Hume: ‘The touchstone of a Humean perception is that everything it is, it also appears to be, *and* everything it appears to be, it also is’ (p. 213). Moreover, Waxman also takes mental transparency to be intimately linked to sensation: ‘To say that perceptions are what they appear to be can only mean, in view of Hume’s theory of belief, that immediate consciousness *feels* them really to be all and only what is met with in (the nature, situation, and relations in) their appearance’ (p. 215).

³⁰ Schmaltz (1996) and Jorgenson (2010) argues that Malabranche follows the Cartesian La Forge in this treatment of consciousness. The passages from *The Search after Truth* refer to Lennon’s and Olscamp’s translation, Cambridge University Press (1997).

³¹ C.f. THN 1.1.1.1 (SBN 1); THN 1.2.6.2 (SBN 66); THN 1.3.14.6 (SBN 158); THN 1.4.2.47 (SBN 212); THN 1.4.7.3 (SBN 265); THN 2.1.2.2 (SBN 277); THN 2.1.11.6 (SBN 318); THN App. 4 (SBN 625); EHU 7.9 (SBN 64); EHU 7.14 (SBN 66); and EHU 7.25 (SBN 71).

³² Admittedly, ‘consciousness’ also carries a connotation of higher-order perception, in accordance with Baxter’s Apperceptual Incorrigeability; indeed we saw that Hume characterises it as such in THN App 20 (SBN 635). The first of Clarke’s (1731, 177n) six definitions is ‘the Reflex Act, by which a Man knows his Thoughts to be his own Thoughts’, and indeed he takes this to be the ‘strict and properest sense of the word’. More recently, Lycan (1996) and Armstrong (1981) take consciousness to consist in an inner sense or perception of our mental states, while Rosenthal (1993) holds that consciousness consists in higher-order *thought* rather than perception. Such a connotation of higher-order perception explains why Hume (or those to whom he attributes such a usage) would characterise consciousness as a ‘reflected thought or perception’ in the Appendix while using it in a pre-philosophical sense (rather than as a technical term). As mentioned earlier, my contention is that Hume employs the term in a loose sense with many different connotations. And indeed, we will see that he is more naturally read as using ‘consciousness’ with respect to its connotation of qualitative character rather than higher-order perception in EHU 7.13 (SBN 66) and THN 1.4.2.7 (SBN 190).

³³ Note that my reconstruction does not follow the order of Hume’s statements in THN 1.4.2.7 (SBN 190), because Hume is arguing backwards in this passage. He begins by stating his conclusion (that the senses cannot deceive us with respect to the situation and relation of our impressions), supports this by stating his intermediate conclusion (that actions and sensations of the mind cannot deceive us), and finishes by giving his initial premise, Qualitative Incorrigeability. I consider the following reconstruction to be faithful to Hume’s argument.

³⁴ I do not claim that Hume uses ‘actions of the mind’ in this way in all other contexts; as with his use of ‘consciousness’ I find it difficult to discern a common thread amongst his myriad usages of the term. Care in terminology, I think it is safe to say, is not Hume’s strongest suit. Note that Hume’s argument here does not require mention of ‘actions of the mind’, but as they are the other class of perceptions that cannot deceive us (also being non-representational), Hume probably decided to mention them.

³⁵ An alternative reading of this sentence is that we cannot be mistaken about perceptions *being perceptions*, stating the impossibility of perceptions to feeling appearing different *from perceptions*. This reading seems unlikely, given that Hume thinks that we mistake perceptions for external objects.

³⁶ This is certainly a controversial claim, but defending it adequately would take me too far afield. Cohon and Owen (1997) argue strongly for the claim that Humean impressions cannot represent, and I draw substantially from their paper in this discussion. As Cohon and Owen note, Hume sometimes slips into speaking as though sensations represented real objects (e.g. EHU 12.9; SBN 152), but this is simply a consequence of his being a philosopher and therefore being seduced by the doctrine of ‘the double existence of perceptions and objects’, which, despite its allure, is ‘monstrous’ (THN 1.4.2.54; SBN 216).

³⁷ Being deceived about our sensations by means of a further idea that misrepresents them would not count as the *senses* deceiving us; it would be the faculty that conveyed the misrepresenting idea that was deceiving us.

³⁸ Of course, we can be (and are) mistaken regarding the ‘situation and relations’ of our sensations; the guilty faculty will deceive us by conveying perceptions which misrepresent the extrinsic properties of our sensations. Hume eventually identifies the imagination as the faculty responsible for our taking sensations to have distinct and continued existence apart from ourselves – unlike the senses, the imagination conveys representational ideas, and therefore is perfectly capable of deceiving us regarding the situation and relations of our impressions.

³⁹ Like Hume, I use ‘power’ and ‘necessary connection’ interchangeably; cf. THN 1.3.14.4 (SBN 157).

⁴⁰ In correspondence, Jonathan Cottrell notes that Hume could simply have examined the case whereby the man is stricken on one side and attempts to lift both his arms at the same time. This is a sensible proposal, but it is not the way that Hume sets up his case, as he compares the stricken man to a man in *perfect* health. Hume probably considered such a thought experiment too contrived for his readers to get a firm grip on their intuitions.

⁴¹ The two perceptions differ with respect to vivacity, but it seems plausible that we can discern their sameness or otherwise with certainty even if one is a fainter version of the other.

⁴² There is a worry here: if, as in the case of the missing shade of blue, we can imagine a novel simple idea of volition, then the argument above does not go through. However, Hume does not believe the case of the missing shade of blue to generalise, noting that ‘this instance is so singular, that it is scarcely worth our observing, and does not merit, that for it alone we should alter our general maxim’ (EHU 2.8; SBN 21). Garrett (1997) and Fogelin (1984) argue that the missing shade of blue is peculiar in that it involves a gradation of closely resembling impressions; this is not the case with regard to the impression of volition, and so it seems unlikely that we could imagine a novel simple idea of volition. Thanks to Don Garrett and Jonathan Cottrell for raising this worry.

⁴³ References to Berkeley are the *Principles of Human Knowledge*, edited by Robinson, Oxford University Press (1996). Arabic numerals refer to paragraph numbers.

⁴⁴ For detailed comments on previous drafts, a tremendous debt of gratitude is owed to Don Garrett, Jonathan Cottrell, Donald Ainslie, Donald Baxter, John Biro, David Chalmers, and John Richardson. For excellent questions and comments, I am also indebted to audiences at the Hume Society Conference 2012, held at the University of Calgary, and the New England Colloquium in Early Modern Philosophy 2012, held at Harvard University.